

ABSTRACT OF THE DISCLOSURE

A system for distributing incoming client requests across multiple servers in a networked client-server computer environment processes all requests as a set that occur within a given time interval and collects information on both the attributes of the requests and the resource capability of the servers to dynamically allocate the requests in a set to the appropriate servers upon the completion of the time interval. Preferably, the system includes a request table to collect at least two requests incoming within a predetermined time interval. A request examiner routine analyzes each collected request with respect to at least one attribute. A system status monitor collects resource capability information of each server in a resource table. An optimization and allocation process distributes collected requests in the request table across the multiple servers upon completion of said time interval based on an optimization of potential pairings of the requests in the request table with the servers in the resource table. The optimization and allocation process preferably analyzes metrics maintained in the request table and resource table as part of a relational database to allocate requests to servers based on a minimization of the metric distance between pairings of requests and servers. Preferably, the request table is part of a dynamic, relational database and a process of statistical inference for ascertaining expected demand patterns involving said the attributes adds predictive information about client requests as part of the request examiner routine.